

Pemetaan Ketimpangan Digital Antarprovinsi di Indonesia Berdasarkan Indikator Akses dan Penggunaan Teknologi Menggunakan Algoritma K-Means

Ulfa Pauziah

Teknik Informatika, FTIK, Universitas Indraprasta PGRI

*Email Korespondensi: pelangi_ulfa@yahoo.com

Sejarah Artikel:

Diterima 14-08-2026
Disetujui 18-08-2026
Diterbitkan 20-08-2026

ABSTRACT

Digital inequality remains a challenge in ensuring equitable access to and use of information and communication technology across regions in Indonesia. This study aimed to map digital inequality among 38 Indonesian provinces in 2024 using the K-Means clustering algorithm. Secondary data obtained from Statistics Indonesia (BPS) were analyzed using three indicators: the percentage of the population accessing the internet, owning or controlling mobile phones, and using computers. The data were standardized using Z-score transformation, while the optimal number of clusters was evaluated using the Elbow Method and Silhouette Score. The results showed that two clusters provided the optimal configuration, with a Silhouette Score of 0.7457. Cluster 0 consisted of 36 provinces, with average internet access, mobile phone ownership or control, and computer use of 70.84%, 69.08%, and 12.09%, respectively. Cluster 1 consisted of Central Papua and Highland Papua, with substantially lower averages of 17.74%, 24.73%, and 2.98%, respectively. Principal Component Analysis further supported the clustering structure, with two principal components explaining 98.89% of the total variance. These findings demonstrate a pronounced multidimensional digital gap and highlight the need for more targeted digital development policies in disadvantaged regions.

Keywords: digital inequality; K-Means; internet access; digital divide; provincial clustering

ABSTRAK

Ketimpangan digital masih menjadi tantangan dalam mewujudkan pemerataan akses dan pemanfaatan teknologi informasi dan komunikasi antarwilayah di Indonesia. Penelitian ini bertujuan memetakan ketimpangan digital pada 38 provinsi di Indonesia tahun 2024 menggunakan algoritma K-Means. Data sekunder yang bersumber dari Badan Pusat Statistik (BPS) dianalisis berdasarkan tiga indikator, yaitu persentase penduduk yang mengakses internet, memiliki atau menguasai telepon seluler, dan menggunakan komputer. Data distandardisasi menggunakan transformasi Z-score, sedangkan jumlah cluster optimal ditentukan menggunakan Elbow Method dan Silhouette Score. Hasil penelitian menunjukkan bahwa dua cluster merupakan konfigurasi optimal dengan Silhouette Score sebesar 0,7457. Cluster 0 terdiri atas 36 provinsi dengan rata-rata akses internet 70,84%, kepemilikan atau penguasaan telepon seluler 69,08%, dan penggunaan komputer 12,09%. Cluster 1 terdiri atas Papua Tengah dan Papua Pegunungan dengan rata-rata masing-masing sebesar 17,74%, 24,73%, dan 2,98%. Principal Component Analysis memperkuat struktur pengelompokan dengan dua komponen utama yang menjelaskan 98,89% variasi data. Temuan menunjukkan adanya kesenjangan digital multidimensional yang menonjol dan perlunya kebijakan

pembangunan digital yang lebih terarah pada wilayah dengan keterbatasan akses dan penggunaan teknologi.

Katakunci: akses internet; ketimpangan digital; K-Means; kesenjangan digital; klasterisasi provinsi

Bagaimana Cara Sitasi Artikel ini:

Pauziah, U. (2026). Pemetaan Ketimpangan Digital Antarprovinsi di Indonesia Berdasarkan Indikator Akses dan Penggunaan Teknologi Menggunakan Algoritma K-Means. Jejak Digital: Jurnal Ilmiah Multidisiplin, 2(5), 9863-9874. <https://doi.org/10.63822/xaz20569>

PENDAHULUAN

Transformasi digital semakin berperan penting dalam berbagai sektor, termasuk ekonomi, pendidikan, tata kelola pemerintahan, komunikasi, dan kehidupan sosial, seiring dengan pesatnya perkembangan Teknologi Informasi dan Komunikasi (TIK). Integrasi internet, telepon seluler, dan komputer telah menjadi bagian penting dalam mendukung partisipasi digital serta mengubah pola aktivitas dan interaksi masyarakat (Hanna, 2010). Meskipun demikian, manfaat transformasi digital belum dirasakan secara merata karena masih terdapat perbedaan dalam akses, kepemilikan, dan pemanfaatan teknologi antarkelompok maupun antarwilayah. Kondisi tersebut membentuk kesenjangan digital (*digital divide*), yaitu ketidaksetaraan dalam memperoleh dan memanfaatkan sumber daya teknologi informasi. Kesenjangan digital tidak hanya berkaitan dengan ketersediaan infrastruktur dan perangkat, tetapi juga berpotensi memperbesar perbedaan kesempatan dalam memperoleh informasi dan berpartisipasi dalam aktivitas sosial maupun ekonomi. Oleh karena itu, pemerataan akses dan pemanfaatan teknologi menjadi aspek penting dalam mewujudkan transformasi digital yang inklusif dan memberikan manfaat secara lebih merata (Liu et al., 2023).

Ketimpangan digital di Indonesia masih terlihat jelas ketika ditinjau dari perbedaan akses dan pemanfaatan teknologi antarwilayah. Data Badan Pusat Statistik (BPS) tahun 2024 menunjukkan adanya variasi yang lebar antarprovinsi pada tiga indikator utama, yaitu akses internet, kepemilikan atau penguasaan telepon seluler, dan penggunaan komputer. Papua Pegunungan, misalnya, mencatat persentase penduduk yang mengakses internet hanya sebesar 6,76%, memiliki atau menguasai telepon seluler sebesar 15,76%, dan menggunakan komputer sebesar 1,16%. Sebaliknya, Kepulauan Riau menunjukkan capaian yang jauh lebih tinggi, dengan persentase akses internet mencapai 89,26% dan kepemilikan atau penguasaan telepon seluler sebesar 83,90%. Perbedaan capaian tersebut menunjukkan bahwa peningkatan transformasi digital pada tingkat nasional belum sepenuhnya diikuti oleh pemerataan akses dan pemanfaatan teknologi antarprovinsi. Sejumlah penelitian juga menunjukkan bahwa kesenjangan digital antarwilayah berkaitan dengan karakteristik sosial dan ekonomi, antara lain pendidikan, pendapatan atau PDRB per kapita, serta kondisi ketenagakerjaan (Kartiasih et al., 2022; Koswara & Koswara, 2025). Dengan demikian, tingginya capaian digital secara nasional belum secara otomatis menggambarkan kondisi setiap wilayah, sehingga diperlukan analisis yang mampu mengidentifikasi pola kesenjangan digital antarprovinsi secara lebih komprehensif.

Pemetaan ketimpangan digital menjadi penting karena statistik agregat pada tingkat nasional belum tentu mampu merepresentasikan disparitas digital yang terjadi antarwilayah. Suatu provinsi, misalnya, dapat menunjukkan tingkat kepemilikan atau penguasaan telepon seluler yang relatif tinggi, tetapi memiliki tingkat akses internet atau penggunaan komputer yang lebih rendah, sehingga penggunaan satu indikator saja belum memadai untuk menggambarkan kondisi digital suatu wilayah secara komprehensif (Novianto et al., 2023; Yudistira & Maulana, 2025). Oleh karena itu, pengukuran ketimpangan digital perlu mempertimbangkan beberapa indikator secara simultan agar perbedaan karakteristik digital antarwilayah dapat diidentifikasi secara lebih representatif (Azizah et al., 2022; Liu et al., 2017). Pendekatan berbasis kemiripan karakteristik memungkinkan setiap wilayah dikelompokkan berdasarkan pola akses dan pemanfaatan teknologi yang dimilikinya. Pemetaan tersebut dapat mengidentifikasi kelompok wilayah dengan karakteristik digital yang relatif serupa sekaligus menunjukkan wilayah yang mengalami kesenjangan paling menonjol (Novianto et al., 2023; Pick et al., 2024). Hasil pemetaan selanjutnya dapat menjadi informasi empiris untuk mendukung penentuan prioritas pemerataan akses dan pemanfaatan Teknologi Informasi dan Komunikasi (TIK) yang lebih sesuai dengan karakteristik masing-masing wilayah.

Sejumlah penelitian telah mengkaji ketimpangan digital dengan menggunakan indikator akses dan pemanfaatan Teknologi Informasi dan Komunikasi (TIK) pada berbagai tingkat wilayah. Di Indonesia, pendekatan *K-Means clustering* telah digunakan untuk mengelompokkan kabupaten/kota berdasarkan kesamaan karakteristik indikator TIK dan menunjukkan adanya disparitas digital antarwilayah, khususnya antara wilayah perkotaan dan kawasan timur Indonesia (Yudistira & Maulana, 2025). Pendekatan pengelompokan juga diterapkan untuk mengidentifikasi pola antarprovinsi berdasarkan penggunaan *e-money* dan tingkat kemiskinan, yang menunjukkan bahwa metode *clustering* dapat digunakan untuk mengungkap heterogenitas karakteristik wilayah yang tidak selalu terlihat melalui statistik agregat (Parera et al., 2025). Dalam konteks internasional, ketimpangan akses TIK juga ditemukan berkaitan dengan kondisi deprivasi dan karakteristik geografis wilayah, sebagaimana ditunjukkan dalam penelitian di Argentina (Loyola & García, 2025). Di sisi lain, perkembangan kajian kesenjangan digital menunjukkan bahwa persoalan tersebut tidak hanya berkaitan dengan ketersediaan akses, tetapi juga mencakup keterampilan digital dan manfaat yang diperoleh masyarakat dari penggunaan teknologi (Shaw, 2025). Temuan-temuan tersebut menunjukkan bahwa analisis berbasis data dan pengelompokan wilayah memiliki potensi untuk mengidentifikasi pola ketimpangan digital secara lebih terstruktur serta menyediakan dasar empiris bagi perumusan kebijakan pemerataan digital yang sesuai dengan karakteristik wilayah.

Meskipun berbagai penelitian telah membahas ketimpangan digital dan menerapkan metode *clustering* untuk mengidentifikasi disparitas antarwilayah, masih terdapat ruang untuk mengembangkan pemetaan yang mengintegrasikan beberapa dimensi akses dan pemanfaatan teknologi secara simultan pada tingkat provinsi. Sebagian kajian cenderung berfokus pada indikator tertentu atau menggunakan analisis deskriptif, sehingga belum sepenuhnya menggambarkan kemiripan dan perbedaan karakteristik digital antarwilayah berdasarkan kombinasi beberapa indikator. Penelitian ini mengisi ruang tersebut dengan memetakan 38 provinsi di Indonesia menggunakan data BPS tahun 2024 berdasarkan tiga indikator yang konsisten pada tingkat penduduk, yaitu akses internet, kepemilikan atau penguasaan telepon seluler, dan penggunaan komputer. Secara metodologis, data distandardisasi sebelum proses *clustering*, sedangkan jumlah cluster dievaluasi menggunakan *Elbow Method* dan *Silhouette Score* untuk memperoleh struktur pengelompokan yang didukung oleh ukuran validitas internal. Selanjutnya, *Principal Component Analysis (PCA)* digunakan untuk merepresentasikan struktur cluster dalam ruang dua dimensi. Dengan demikian, kebaruan penelitian terletak pada integrasi indikator digital terkini dan prosedur evaluasi cluster yang sistematis untuk menghasilkan pemetaan ketimpangan digital antarprovinsi yang lebih terukur.

Berdasarkan permasalahan dan kesenjangan penelitian tersebut, penelitian ini bertujuan untuk memetakan ketimpangan digital antarprovinsi di Indonesia menggunakan algoritma *K-Means* berdasarkan tiga indikator, yaitu akses internet, kepemilikan atau penguasaan telepon seluler, dan penggunaan komputer. Penelitian ini diharapkan memberikan beberapa kontribusi. Pertama, menghasilkan pengelompokan provinsi berdasarkan kemiripan karakteristik akses dan pemanfaatan teknologi digital. Kedua, mengidentifikasi kelompok wilayah yang menunjukkan kesenjangan digital paling menonjol dibandingkan kelompok provinsi lainnya. Ketiga, menyediakan pemetaan berbasis data yang dapat digunakan sebagai informasi empiris dalam mengevaluasi pemerataan transformasi digital antarwilayah. Keempat, penelitian ini menerapkan kerangka analisis yang mengintegrasikan standardisasi data, *K-Means*, evaluasi jumlah cluster menggunakan *Elbow Method* dan *Silhouette Score*, serta *PCA* untuk mendukung interpretasi dan visualisasi struktur pengelompokan. Hasil penelitian diharapkan dapat memberikan gambaran yang lebih terstruktur mengenai disparitas digital antarprovinsi serta menjadi masukan bagi

penentuan prioritas pemerataan akses dan pemanfaatan TIK di Indonesia.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksploratif dengan teknik *clustering* untuk memetakan ketimpangan digital antarprovinsi di Indonesia berdasarkan kemiripan karakteristik penggunaan teknologi informasi dan komunikasi. Metode *clustering* yang digunakan adalah algoritma K-Means, sedangkan Principal Component Analysis (PCA) digunakan sebagai teknik pendukung untuk memvisualisasikan struktur kelompok yang terbentuk. Unit analisis penelitian mencakup 38 provinsi di Indonesia dengan menggunakan data tahun 2024.

Sumber dan Variabel Data

Data yang digunakan merupakan data sekunder yang bersumber dari Badan Pusat Statistik (BPS) tahun 2024. Pemilihan data sekunder dilakukan karena BPS menyediakan indikator teknologi informasi dan komunikasi yang terstandarisasi dan mencakup seluruh provinsi di Indonesia. Penelitian menggunakan nilai gabungan perkotaan dan perdesaan sehingga setiap observasi menggambarkan kondisi provinsi secara keseluruhan.

Tiga indikator digunakan sebagai variabel pembentuk cluster, yaitu: (1) persentase penduduk berumur lima tahun ke atas yang mengakses internet selama tiga bulan terakhir (X_1); (2) persentase penduduk yang memiliki atau menguasai telepon seluler (X_2); dan (3) persentase penduduk yang menggunakan komputer (X_3). Ketiga indikator dipilih untuk merepresentasikan dimensi akses dan penggunaan teknologi digital masyarakat.

Prapemrosesan dan Standardisasi Data

Tahap awal analisis dilakukan dengan mengintegrasikan ketiga dataset berdasarkan nama provinsi. Agregat nasional Indonesia tidak dimasukkan sebagai observasi karena unit analisis penelitian adalah provinsi, sehingga dataset akhir terdiri atas 38 observasi. Selanjutnya dilakukan pemeriksaan tipe data, kesesuaian nama provinsi, serta missing value sebelum proses clustering.

Meskipun ketiga indikator dinyatakan dalam persentase, distribusi dan rentang nilainya berbeda. Oleh karena itu, standardisasi dilakukan menggunakan Z-score agar setiap variabel memberikan kontribusi yang sebanding terhadap perhitungan jarak pada K-Means. Standardisasi dirumuskan sebagai:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$$

Dengan z_{ij} merupakan nilai terstandarisasi provinsi ke- i pada variabel ke- j , x_{ij} merupakan nilai asli, $\bar{x}_j$ merupakan rata-rata variabel ke- j , dan s_j merupakan standar deviasi variabel ke- j .

Penentuan Jumlah Cluster Optimal

Jumlah cluster tidak ditetapkan secara subjektif, tetapi dievaluasi untuk $K = 2$ hingga $K = 6$ menggunakan **Elbow Method** dan **Silhouette Score**. *Elbow Method* mengevaluasi perubahan *within-cluster sum of squares* atau *inertia* ketika jumlah cluster bertambah. Titik ketika penurunan *inertia* mulai melandai digunakan sebagai salah satu pertimbangan dalam menentukan jumlah cluster.

Untuk memperkuat pemilihan jumlah cluster, penelitian menggunakan *Silhouette Score* yang mengukur tingkat kohesi objek terhadap cluster-nya sekaligus separasi terhadap cluster lainnya. Nilai *silhouette* untuk observasi dihitung sebagai:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}$$

Dengan $a(i)$ merupakan rata-rata jarak observasi i terhadap observasi lain dalam cluster yang sama, sedangkan $b(i)$ merupakan rata-rata jarak minimum observasi i terhadap cluster terdekat lainnya. Nilai yang semakin mendekati 1 menunjukkan struktur cluster yang semakin baik.

Klasterisasi Menggunakan K-Means

Setelah jumlah cluster optimal ditentukan, proses pengelompokan dilakukan menggunakan algoritma K-Means. Algoritma ini mengelompokkan provinsi berdasarkan kedekatan terhadap pusat cluster (centroid) dengan meminimalkan variasi dalam masing-masing cluster. Secara matematis, fungsi objektif K-Means dapat dinyatakan sebagai:

$$J = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - \mu_k\|^2$$

Dengan K merupakan jumlah cluster, C_k adalah cluster ke- k , x_i merupakan observasi provinsi, dan μ_k merupakan centroid cluster ke- k . Proses iterasi dilakukan sampai tidak terjadi perubahan berarti pada keanggotaan cluster dan posisi centroid. Untuk menjaga reproduktibilitas analisis, proses K-Means menggunakan `random_state=42` dan `n_init=10`.

Visualisasi Menggunakan PCA

Principal Component Analysis (PCA) digunakan setelah proses clustering untuk membantu visualisasi struktur kelompok dalam ruang dua dimensi. PCA mentransformasikan tiga variabel terstandarisasi menjadi dua komponen utama, yaitu PC1 dan PC2, dengan mempertahankan sebanyak mungkin variasi data asli. PCA pada penelitian ini berfungsi sebagai alat visualisasi, bukan sebagai variabel input pengganti pada proses K-Means. Dengan demikian, hasil clustering tetap diperoleh berdasarkan tiga indikator digital yang telah distandarisasi.

HASIL DAN PEMBAHASAN

1. Statistik Deskriptif Indikator Ketimpangan Digital

Analisis deskriptif dilakukan untuk memberikan gambaran awal mengenai variasi tingkat akses dan penggunaan teknologi digital antarprovinsi di Indonesia. Tiga indikator yang dianalisis meliputi persentase penduduk yang mengakses internet, persentase penduduk yang memiliki atau menguasai telepon seluler, dan persentase penduduk yang menggunakan komputer. Perbedaan nilai antarprovinsi pada ketiga indikator tersebut menunjukkan adanya heterogenitas karakteristik digital yang menjadi dasar dalam proses pengelompokan menggunakan algoritma K-Means.

Tabel 1. Statistik Deskriptif Indikator Digital Antarprovinsi di Indonesia Tahun 2024

Indikator	N	Rata-rata	Std. Deviasi	Minimum	Maksimum
Akses Internet	38	68,04	14,83	6,76	89,26
Kepemilikan/Penguasaan Telepon Seluler	38	66,75	12,26	15,76	83,90
Penggunaan Komputer	38	11,61	3,82	1,1	1,16

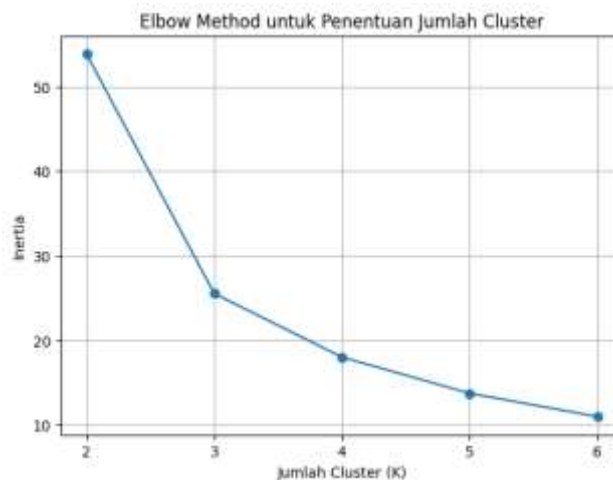
Sumber: Data BPS 2024, diolah menggunakan Python.

Hasil analisis deskriptif menunjukkan adanya variasi capaian indikator digital yang cukup besar antarprovinsi di Indonesia. Rata-rata persentase penduduk yang mengakses internet pada tahun 2024 mencapai 68,04%, dengan standar deviasi sebesar 14,83. Nilai akses internet berkisar antara 6,76% hingga 89,26%, menunjukkan adanya rentang yang lebar antara provinsi dengan tingkat akses internet terendah dan tertinggi. Kondisi tersebut memberikan indikasi awal bahwa akses terhadap konektivitas digital belum tersebar secara merata antarprovinsi.

Pada indikator kepemilikan atau penguasaan telepon seluler, rata-rata capaian sebesar 66,75%, dengan standar deviasi 12,26. Nilai minimum tercatat sebesar 15,76%, sedangkan nilai maksimum mencapai 83,90%. Sementara itu, penggunaan komputer memiliki rata-rata yang jauh lebih rendah, yaitu 11,61%, dengan standar deviasi 3,82 serta rentang antara 1,16% hingga 23,47%. Perbedaan rentang dan penyebaran pada ketiga indikator tersebut memperlihatkan heterogenitas karakteristik digital antarprovinsi dan menjadi dasar empiris untuk melakukan pengelompokan menggunakan algoritma K-Means.

2. Penentuan Jumlah Cluster Optimal

Sebelum proses K-Means dilakukan, jumlah cluster optimal dievaluasi menggunakan *Elbow Method* dan *Silhouette Score*. Pengujian dilakukan untuk K=2 sampai K=6. *Elbow Method* digunakan untuk mengamati perubahan nilai *inertia*, sedangkan *Silhouette Score* digunakan untuk mengevaluasi tingkat kohesi objek dalam cluster dan separasinya terhadap cluster lain.



Gambar 1. Elbow Method untuk Penentuan Jumlah Cluster

Grafik *Elbow Method* menunjukkan penurunan *inertia* yang tajam pada jumlah cluster awal dan kemudian mengalami penurunan yang semakin kecil seiring bertambahnya jumlah cluster. Untuk memperoleh dasar pemilihan yang lebih objektif, hasil tersebut dibandingkan dengan *Silhouette Score*

sebagaimana disajikan pada Tabel 2.

Tabel 2. Evaluasi Jumlah Cluster Berdasarkan Silhouette Score

Jumlah Cluster (K)	Silhouette Score
2	0,7457
3	0,5989
4	0,3444
5	0,3437
6	0,3343

Sumber: Hasil pengolahan data menggunakan Python.

Nilai *Silhouette Score* tertinggi diperoleh pada K=2 sebesar 0,7457, diikuti K=3 sebesar 0,5989. Ketika jumlah cluster ditingkatkan menjadi empat hingga enam, nilai *Silhouette Score* menurun menjadi sekitar 0,33–0,34. Hasil tersebut menunjukkan bahwa pembentukan dua cluster memberikan struktur pengelompokan paling jelas dibandingkan konfigurasi lainnya. Oleh karena itu, K=2 ditetapkan sebagai jumlah cluster optimal dan digunakan dalam analisis K-Means selanjutnya.

3. Hasil Klasterisasi K-Means

Penerapan algoritma K-Means dengan dua cluster menghasilkan struktur pengelompokan yang sangat tidak seimbang dari sisi jumlah anggota, tetapi memiliki perbedaan karakteristik digital yang kuat. Cluster 0 terdiri atas 36 provinsi, sedangkan Cluster 1 hanya terdiri atas dua provinsi. Interpretasi karakteristik kedua cluster dilakukan berdasarkan nilai rata-rata setiap indikator pada masing-masing cluster atau *centroid*, sebagaimana ditunjukkan pada Tabel 3.

Tabel 3. Karakteristik Centroid Hasil Klasterisasi K-Means

Cluster	Akses Internet (%)	Telepon Seluler (%)	Penggunaan Komputer (%)	Jumlah Provinsi
Cluster 0	70,84	69,08	12,09	36
Cluster 1	17,74	24,73	2,98	2

Sumber: Hasil pengolahan data menggunakan Python.

Cluster 0 memiliki centroid akses internet sebesar 70,84%, kepemilikan/penguasaan telepon seluler sebesar 69,08%, dan penggunaan komputer sebesar 12,09%. Sebaliknya, Cluster 1 hanya memiliki rata-rata akses internet 17,74%, telepon seluler 24,73%, dan penggunaan komputer 2,98%. Selisih centroid antara kedua cluster mencapai 53,10 poin persentase pada akses internet, 44,35 poin persentase pada telepon seluler, dan 9,11 poin persentase pada penggunaan komputer.

Perbedaan tersebut menunjukkan bahwa pemisahan cluster tidak disebabkan oleh satu indikator saja. Cluster 1 secara konsisten memiliki nilai lebih rendah pada seluruh indikator yang digunakan. Oleh karena itu, Cluster 0 dapat diinterpretasikan sebagai kelompok utama dengan capaian digital relatif lebih tinggi, sedangkan Cluster 1 merepresentasikan kelompok dengan kesenjangan digital tinggi. Penggunaan istilah relatif penting karena K-Means membentuk kelompok berdasarkan kemiripan karakteristik data dan bukan berdasarkan suatu ambang normatif mengenai tingkat kemajuan digital.

4. Provinsi dengan Kesenjangan Digital Tinggi

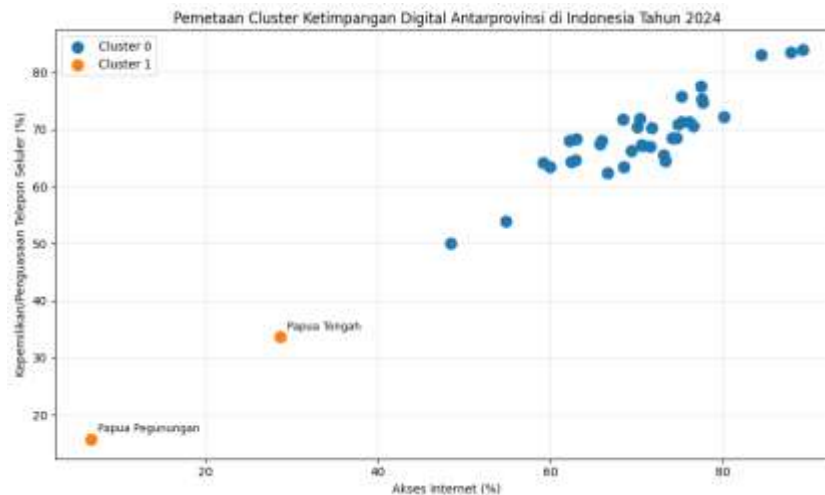
Hasil pengelompokan menunjukkan bahwa Cluster 1 hanya terdiri atas Papua Tengah dan Papua Pegunungan. Nilai masing-masing indikator kedua provinsi disajikan pada Tabel 4.

Tabel 4. Provinsi pada Cluster Kesenjangan Digital Tinggi

Provinsi	Akses Internet (%)	Telepon Seluler (%)	Penggunaan Komputer (%)
Papua Tengah	28,71	33,70	4,81
Papua Pegunungan	6,76	15,76	1,16

Sumber: Data BPS 2024, diolah menggunakan Python.

Papua Tengah menunjukkan akses internet sebesar 28,71%, kepemilikan atau penguasaan telepon seluler sebesar 33,70%, dan penggunaan komputer sebesar 4,81%. Kondisi yang lebih ekstrem ditemukan pada Papua Pegunungan, dengan akses internet hanya 6,76%, telepon seluler 15,76%, dan penggunaan komputer 1,16%. Seluruh nilai tersebut berada jauh di bawah centroid Cluster 0.



Gambar 2. Pemetaan Cluster Ketimpangan Digital Antarprovinsi di Indonesia Tahun 2024

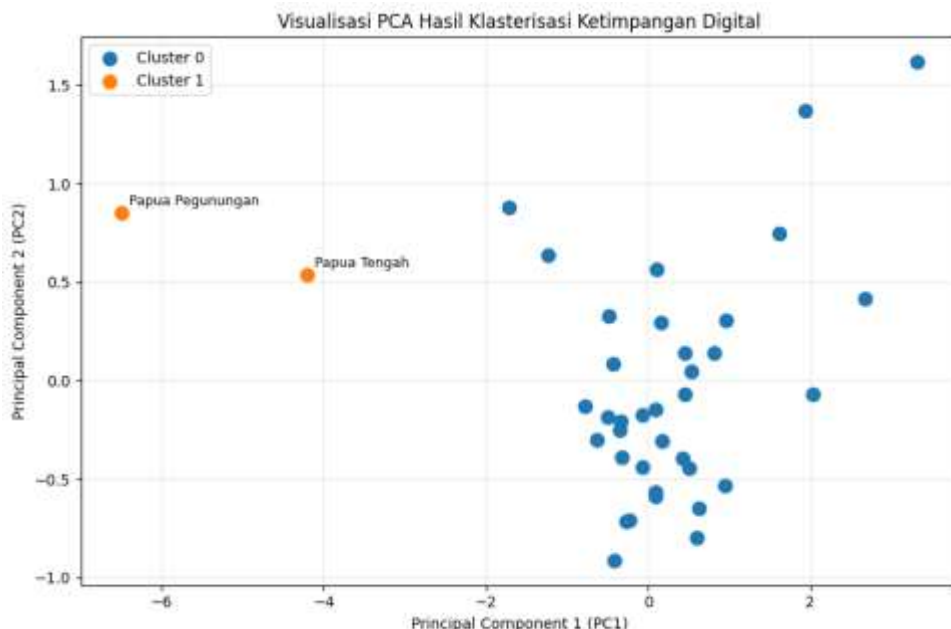
Visualisasi cluster memperlihatkan bahwa sebagian besar provinsi terkonsentrasi pada area dengan tingkat akses internet dan penguasaan telepon seluler yang relatif lebih tinggi. Sebaliknya, Papua Tengah dan Papua Pegunungan berada pada posisi yang terpisah dari kelompok utama. Papua Pegunungan menunjukkan posisi paling ekstrem, sedangkan Papua Tengah berada di antara Papua Pegunungan dan kelompok utama tetapi tetap cukup jauh sehingga tergabung dalam cluster yang sama dengan Papua Pegunungan.

Temuan tersebut menjadi aspek penting dalam pemetaan ketimpangan digital Indonesia. Hasil clustering tidak sekadar menunjukkan adanya variasi antarprovinsi, tetapi mengidentifikasi adanya dua provinsi dengan profil digital yang secara multivariat berbeda dari mayoritas provinsi lainnya. Dengan demikian, kesenjangan yang teridentifikasi bukan hanya persoalan rendahnya akses internet, melainkan juga rendahnya penguasaan telepon seluler dan penggunaan komputer secara simultan.

5. Visualisasi Struktur Cluster Menggunakan PCA

Untuk mengevaluasi representasi visual struktur data tiga dimensi, hasil clustering divisualisasikan menggunakan Principal Component Analysis (PCA). PCA mereduksi tiga indikator terstandarisasi menjadi dua komponen utama tanpa digunakan untuk menentukan keanggotaan cluster.

Hasil PCA menunjukkan bahwa PC1 menjelaskan 87,48% variasi data, sedangkan PC2 menjelaskan 11,41%. Secara kumulatif, kedua komponen mampu menjelaskan 98,89% variasi data awal. Persentase tersebut menunjukkan bahwa hampir seluruh informasi variasi dari tiga indikator dapat direpresentasikan dalam ruang dua dimensi.



Gambar 3. Visualisasi PCA Hasil Klasterisasi Ketimpangan Digital Antarprovinsi

Visualisasi PCA memperlihatkan pemisahan yang konsisten dengan hasil K-Means. Mayoritas provinsi berada dalam kelompok yang relatif berdekatan, sedangkan Papua Tengah dan Papua Pegunungan terpisah secara jelas, terutama sepanjang PC1. Posisi Papua Pegunungan yang berada paling jauh dari konsentrasi mayoritas provinsi menunjukkan profil indikator digital yang sangat berbeda. Papua Tengah juga berada pada arah yang sama, meskipun jaraknya terhadap kelompok utama tidak sebesar Papua Pegunungan.

Konsistensi antara hasil K-Means, Silhouette Score sebesar 0,7457, dan visualisasi PCA yang mempertahankan 98,89% variasi data memberikan dukungan terhadap struktur dua cluster yang diperoleh. Dengan demikian, pemisahan Papua Tengah dan Papua Pegunungan tidak hanya terlihat pada nilai individual setiap indikator, tetapi juga tetap muncul ketika ketiga indikator dianalisis secara simultan dalam ruang komponen utama.

6. Pembahasan Ketimpangan Digital Antarprovinsi

Secara keseluruhan, hasil penelitian menunjukkan bahwa ketimpangan digital antarprovinsi di Indonesia memiliki pola yang cukup spesifik. Sebanyak 36 dari 38 provinsi berada dalam satu kelompok utama, sementara Papua Tengah dan Papua Pegunungan membentuk kelompok yang berbeda secara

signifikan. Struktur ini menunjukkan bahwa ketimpangan digital yang teridentifikasi melalui tiga indikator tidak tersebar secara seragam di seluruh provinsi, tetapi terdapat wilayah tertentu dengan tingkat keterbatasan digital yang jauh lebih menonjol.

Temuan juga memperlihatkan bahwa akses internet merupakan indikator dengan variasi absolut terbesar, dengan rentang antara 6,76% hingga 89,26%. Namun, rendahnya capaian pada wilayah Cluster 1 tidak terbatas pada akses internet. Kedua provinsi juga menunjukkan tingkat kepemilikan atau penguasaan telepon seluler serta penggunaan komputer yang rendah. Pola tersebut mengindikasikan bahwa kesenjangan digital yang ditemukan bersifat multidimensional. Dengan kata lain, peningkatan pemerataan digital perlu mempertimbangkan konektivitas sekaligus ketersediaan dan penggunaan perangkat digital.

Hasil ini memiliki implikasi penting bagi kebijakan pemerataan transformasi digital. Pendekatan kebijakan yang seragam untuk seluruh provinsi berpotensi kurang efektif karena karakteristik kesenjangan tidak sama. Provinsi yang berada dalam kelompok utama masih dapat memiliki persoalan internal yang tidak ditangkap oleh pengelompokan dua cluster, sedangkan Papua Tengah dan Papua Pegunungan menunjukkan kebutuhan yang lebih mendasar pada ketiga indikator sekaligus. Oleh sebab itu, hasil clustering dapat digunakan sebagai pemetaan awal untuk membantu mengidentifikasi wilayah yang membutuhkan perhatian lebih besar dalam pengembangan akses dan pemanfaatan teknologi digital.

KESIMPULAN

Penelitian ini berhasil memetakan ketimpangan digital antarprovinsi di Indonesia tahun 2024 menggunakan algoritma K-Means berdasarkan tiga indikator, yaitu akses internet, kepemilikan atau penguasaan telepon seluler, dan penggunaan komputer. Hasil evaluasi jumlah cluster menunjukkan bahwa $K=2$ merupakan konfigurasi optimal dengan Silhouette Score tertinggi sebesar 0,7457, lebih tinggi dibandingkan $K=3$ hingga $K=6$. Hasil klasterisasi membentuk dua kelompok yang memiliki karakteristik digital berbeda secara jelas. Cluster 0 terdiri atas 36 provinsi dengan rata-rata akses internet sebesar 70,84%, kepemilikan atau penguasaan telepon seluler 69,08%, dan penggunaan komputer 12,09%. Sementara itu, Cluster 1 terdiri atas Papua Tengah dan Papua Pegunungan, dengan rata-rata masing-masing indikator hanya sebesar 17,74%, 24,73%, dan 2,98%. Hasil PCA turut memperkuat struktur pengelompokan tersebut, dengan dua komponen utama mampu menjelaskan 98,89% variasi data. Temuan ini menunjukkan bahwa ketimpangan digital antarprovinsi tidak hanya berkaitan dengan konektivitas internet, tetapi juga mencakup penguasaan dan penggunaan perangkat digital. Oleh karena itu, pemerataan transformasi digital memerlukan pendekatan kebijakan yang lebih terarah, terutama pada wilayah dengan keterbatasan multidimensional seperti Papua Tengah dan Papua Pegunungan. Penelitian selanjutnya dapat memperluas indikator dengan memasukkan kualitas jaringan, keterjangkauan layanan, literasi digital, kondisi sosial ekonomi, dan infrastruktur telekomunikasi serta menggunakan data beberapa tahun untuk menganalisis dinamika ketimpangan digital antarwilayah.

DAFTAR PUSTAKA

Azizah, F., Pramesti, W., & Fitriani, F. (2022). Analisis Education Mapping Terkait Pengelompokan Kesenjangan Pembangunan Pendidikan Menurut Provinsi di Indonesia. *MUST: Journal of Mathematics Education, Science and Technology*, 7(2), 130–138. <https://doi.org/10.30651/must.v7i2.11097>

- Hanna, N. K. (2010). *An ICT-Transformed Government and Society* (pp. 1–25). Springer, New York, NY. https://doi.org/10.1007/978-1-4419-1506-1_1
- Kartiasih, F., Nachrowi, N. D., Wisana, I. D. G. K., & Handayani, D. (2022). Inequalities of Indonesia's regional digital development and its association with socioeconomic characteristics: a spatial and multivariate analysis. *Information Technology for Development*, 1–30. <https://doi.org/10.1080/02681102.2022.2110556>
- Koswara, A., & Koswara, A. (2025). *Digital Inequality: E-Commerce Access and Its Implications for Underdeveloped Regions in Indonesia*. 2(1), 1–14. <https://doi.org/10.65119/jspds.v2i1.20>
- Liu, H., Chen, J.-C., Dy, J. G., & Fu, Y. (2023). Transforming Complex Problems Into K-Means Solutions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45, 9149–9168. <https://doi.org/10.1109/TPAMI.2023.3237667>
- Liu, H., Fang, C., & Sun, S. (2017). Digital inequality in provincial China. *Environment and Planning A*, 49(10), 2179–2182. <https://doi.org/10.1177/0308518X17711946>
- Loyola, V. F., & García, J. R. (2025). Spatial Patterns of ICT Access in Argentine Households: Regional and Departmental Analysis (2022). *Urban Science*. <https://doi.org/10.3390/urbansci9120537>
- Novianto, M. D., Sumantri, G., & Prihastuti, P. P. (2023). Clustering Provinsi di Indonesia Berdasarkan Indeks Pembangunan Teknologi Informasi dan Komunikasi Menggunakan K-Medoids. *Emerging Statistics and Data Science Journal*, 1(3), 320–331. <https://doi.org/10.20885/esds.vol1.iss.3.art39>
- Parera, N. M., Dewi, C., & Purnomo, H. D. (2025). *Mapping The Digital Divide: A Spatial Clustering Analysis of E-Money Adoption and Regional Poverty in Indonesia*. 1–4. <https://doi.org/10.1109/icast68191.2025.11300019>
- Pick, J. B., Ren, F., & Sarkar, A. (2024). Digital Inequalities in China in 2020: Spatial and Multivariate Analysis. *Applied Sciences*, 14(13), 5385. <https://doi.org/10.3390/app14135385>
- Shaw, K. (2025). *Levels of digital divide: A Review Literature*. <https://doi.org/10.5281/zenodo.17227871>
- Yudistira, M. R., & Maulana, R. I. (2025). Pengelompokan Kabupaten/Kota di Indonesia Berdasarkan Indikator Teknologi Informasi dan Komunikasi Menggunakan Metode K-Means. *Prosiding Seminar Nasional Official Statistics*. <https://doi.org/10.34123/semnasoffstat.v2025i1.2424>